Technical Report 1220

Training for Efficient, Durable, and Flexible Performance in the Military

Alice F. Healy
Erica L. Wohldmann
James A. Kole
Vivian I. Schneider
Kathleen M. Shea
Lyle E. Bourne, Jr.
Center for Research on Training
University of Colorado

January 2008

United States Army Research Institute for the Behavioral and Social Sciences

U.S. Army Research Institute
for the Behavioral and Social Sciences

A Directorate of the Department of the Army
Deputy Chief of Staff, G1

Authorized and approved for distribution:

Michelle Sams

MICHELLE SAMS, Ph.D.
Director

Research accomplished under contract
for the Department of the Army

Center for Research on Training
University of Colorado

Technical review by:

David P. Costanza, The George Washington University
Jonathan Kaplan, U. S. Army Research Institute

NOTICES

DISTRIBUTION: Primary distribution of this Technical Report has been made by ARI. Please address correspondence concerning distribution of reports to: U.S. Army Research Institute for the Behavioral and Social Sciences, Attn: DAPE-ARI-MS, 2511 Jefferson Davis Highway, Arlington, Virginia 22202-3926.

FINAL DISPOSITION: This Technical Report may be destroyed when it is no longer needed. Please do not return it to the U.S. Army Research Institute for the Behavioral and Social Sciences.

NOTE: The findings in this Technical Report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents

REPORT DOCUMENTATION PAGE

1. REPORT DATE (dd-mm-yy) January 2008	2. REPORT TYPE Final	3. DATES COVERED (from. . . to) September 2002 – September 2007
4. TITLE AND SUBTITLE Training for Efficient, Durable, and Flexible Performance in the Military		5a. CONTRACT OR GRANT NUMBER DASW01-03-K-0002 5b. PROGRAM ELEMENT NUMBER 611102
6. AUTHOR(S) Alice F. Healy, Erica L. Wohldmann, James A. Kole, Vivian I. Schneider, Kathleen M. Shea, Lyle E. Bourne, Jr. (University of Colorado)		5c. PROJECT NUMBER B74F 5d. TASK NUMBER 5e. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Regents of the University of Colorado Randall Draper Office of Contracts & Grants University Boulder, CO 80309-0572		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) U. S. Army Research Institute for the Behavioral & Social Sciences 2511 Jefferson Davis Highway Arlington, VA 22202-3926		10. MONITOR ACRONYM ARI 11. MONITOR REPORT NUMBER Technical Report 1220

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

Contractor Officer's Representative and Subject Matter POC: Dr. Paul A. Gade

14. ABSTRACT *(Maximum 200 words)*:

Progress has been made on the following topics concerning training for efficient, durable, and flexible performance in the military: (a) dealing with information flow, (b) factors promoting adaptive and flexible performance, and (c) coping with dynamic environments and changing task demands.

15. SUBJECT TERMS

Training, transfer, long-term retention, knowledge, and skills

SECURITY CLASSIFICATION OF			19. LIMITATION OF ABSTRACT	20. NUMBER OF PAGES	21. RESPONSIBLE PERSON
16. REPORT Unclassified	17. ABSTRACT Unclassified	18. THIS PAGE Unclassified	Unlimited	31	Ellen Kinzer Technical Publication Specialist 703-602-8047

Standard Form 298

Technical Report 1220

Training for Efficient, Durable, and Flexible Performance in the Military

Alice F. Healy
Erica L. Wohldmann
James A. Kole
Vivian I. Schneider
Kathleen M. Shea
Lyle E. Bourne, Jr.
Center for Research on Training
University of Colorado

Basic Research Unit
Paul A. Gade, Chief

U.S. Army Research Institute for the Behavioral and Social Sciences
2511 Jefferson Davis Highway, Arlington, Virginia 22202-3926

January 2008

Army Project Number 0611102B74F	**Personnel, Performance and Training**

TRAINING FOR EFFICIENT, DURABLE, AND FLEXIBLE PERFORMANCE IN THE MILITARY

EXECUTIVE SUMMARY

Research Requirement:

To optimize performance in the military, training should be efficient, durable, and flexible. Efficiency is essential because of the high costs of training. Military training also must be durable to ensure long-term retention of the trained knowledge and skills for later success in the field. But durable training will be insufficient if the learned knowledge and skills cannot be applied to situations different from those encountered during training. Because training can rarely capture the full set of circumstances under which tasks are subsequently encountered, another important goal for training is transfer or flexibility.

This report reviews research conducted at the University of Colorado on training for efficient, durable, and flexible performance in the military, with the support of ARI Contract DASW01-03-K-0002.

Procedures:

Five separate lines of research contribute to this report. The first three demonstrate a high degree of specificity of learning. We identified certain circumstances that lead to remarkable durability of what has been learned; yet those same conditions yield very poor flexibility, or the ability to generalize learning to new situations or contexts. Empirical findings are presented illustrating specificity and summarizing our theoretical explanations for the particular tasks we investigated. We propose a general theoretical framework that can account for the high degree of specificity obtained in these studies and also enables us to predict when learning will be generalizable rather than specific. In addition, in support of our theoretical framework, results from two other lines of research are summarized demonstrating situations showing robust transfer of learning.

Findings:

The research summarized in this report might be used to drive applied research. To illustrate this potential symbiosis between basic and applied research, we give two brief examples. First, our research has demonstrated a high degree of specificity from training to subsequent application. In fact, we have shown that training is specific even to the length of messages that need to be understood and executed. Test performance was best following training with all possible message lengths. Second, we have shown that prior knowledge can be used to enhance the learning of spatial position information. More generally, the results from all five lines of research summarized here support the working hypothesis that there is specificity (limited transfer) for tasks based primarily on procedural information, or skill, whereas there is generality (robust transfer) for tasks based primarily on declarative information, or facts. Thus, these studies provide evidence that for skill learning, retention is high but transfer is low; in

contrast, for fact learning, retention is low but transfer is high. This working hypothesis is an expanded version of the procedural reinstatement principle. Our research has validated this principle along with several other training principles, which collectively form the theoretical framework for our studies.

Utilization and Dissemination of Findings:

Our research findings have crucial implications for military training because instructors may assume that teaching a particular task through a limited number of examples will generalize fully to an entire domain even when the examples differ in a fundamental respect (e.g., length) from the test situations. However, our findings imply that to be effective, training should incorporate a full range of examples on critical task dimensions. Although the tasks used in our research are often components of military tasks, they are not the real military tasks currently being trained in the Army. We hope that applied research units are interested in testing whether the principles we have developed would apply to such real tasks and whether the methods we hope to develop for overcoming the problem of training specificity could be adapted to improve military training.

TRAINING FOR EFFICIENT, DURABLE, AND FLEXIBLE PERFORMANCE IN THE MILITARY

CONTENTS

Page

LIST OF FIGURES

Introduction

There are three aspects of training we should consider when we are trying to optimize human performance. First is the efficiency of training. Because of the high costs of training, we certainly want to be sure that training is accomplished as quickly as possible without sacrificing the level of achievement. However, optimizing training time should not be the only, or even the most important, goal. If individuals have successfully learned how to perform a task during training but then forget how to perform it sometime later, the training has clearly been ineffective. Passing a test at the end of training does not guarantee later success in the field. Long-term retention of the trained knowledge and skills is essential. The second aspect is that learning must be durable as well as efficient. But even durable training will not be sufficient if the learned knowledge and skills cannot be applied to situations different from those encountered during training. Because training can rarely capture the full set of circumstances under which tasks are subsequently encountered, the third important goal for training is transfer or flexibility. Optimal training, thus, should be efficient, durable, and flexible. This report reviews research we have been conducting at the University of Colorado on training for efficient, durable, and flexible performance in the military, with the support of ARI Contract DASW01-03-K-0002. Over the course of this contract, we completed a total of 23 experiments addressed to the following three general topics: (a) dealing with information flow, (b) factors promoting adaptive and flexible performance, and (c) coping with dynamic environments and changing task demands. These experiments, which have illuminated factors optimizing the efficiency, durability, and flexibility of training, have been described in previously submitted annual reports; a detailed description of them will not be provided here. Instead, we concentrate in this report on general conclusions that can be drawn from this research; towards this end, we summarize here a selected subset of the completed experiments that focus on the durability and specificity of training.

This report is centered on five separate lines of research. The first three lines demonstrate a high degree of specificity of learning. We have identified certain circumstances leading to remarkable durability of what has been learned. But those same conditions yield very poor flexibility, or the ability to generalize learning to new situations or contexts. We will present empirical findings illustrating specificity and briefly summarize our theoretical explanations of them for the particular tasks we investigated. We will follow this review by summarizing results from two other lines of research demonstrating situations showing powerful transfer of learning, again briefly summarizing our theoretical explanations for the particular tasks involved. In addition to these restricted theoretical explanations, we propose here a more comprehensive theoretical framework that can account both for the weak flexibility obtained in the first set of three studies and for the strong flexibility found in the second set of two studies. This framework enables us to predict when learning will be generalizable rather than specific. We will conclude with a short discussion of the relevance of this work to military training.

There have been at least two principles proposed in the literature that are consistent with the specificity of training we observed in the first three lines of investigation. By the *encoding specificity principle* (Tulving & Thomson, 1973), retrieval is successful to the extent that the encoding cues and operations correspond with those available at retrieval. By the *transfer appropriate processing principle* (McDaniel, Friedman, & Bourne, 1978; Morris, Bransford, & Franks, 1977; Roediger, Weldon, & Challis, 1989), performance depends more on the

correspondence between the processing occurring during acquisition and that occurring during testing than on the level of processing during acquisition alone. These two principles were formulated to account for performance in list learning and memory tasks.

Earlier, as a consequence of our previous research on the long-term retention and transfer of knowledge and skills, we proposed a related *procedural reinstatement principle* (Healy & Bourne, 1995; Healy, Wohldmann, & Bourne, 2005). This principle, which is based in part on encoding specificity and transfer appropriate processing, also makes use of the distinction put forth by Anderson (1983) between procedural information (knowing how to do something) and declarative information (knowing that something is the case). Specifically, according to the procedural reinstatement principle, procedural information is more durable than declarative information. However, durable performance lacks generality because, as argued by Kolers and Roediger (1984), performance at test is optimal only when the procedures acquired during training are duplicated during testing.

These principles are consistent with a number of classic and contemporary models, according to which transfer occurs only when there is a match in elements between the activities occurring during training and testing. Thorndike (1906) was the first to propose a theory of "identical elements," the elements in his theory being stimulus-response associations. Subsequent related models were proposed by Singley and Anderson (1989) and by Rickard and Bourne (1995). For Singley and Anderson, the elements were production (condition-action) rules, and for Rickard and Bourne, who were primarily concerned with mental arithmetic, the elements were abstract representations of numbers and arithmetic operations (see Rickard, 2005, for a revised version of this model).

The high degree of specificity of transfer implied by these principles and models needs to be tempered with an acknowledgment of the generality of transfer found in many other studies. Indeed the specificity of training we have shown in the first three lines of investigation is actually quite surprising given the findings from many other investigations reported in the literature (e.g., Harlow, 1949). To capture all available results, we need to invoke a critical distinction between specificity and generality of transfer as well as the one between procedural and declarative knowledge. An expansion of the procedural reinstatement principle provides a working hypothesis (Healy, 2007) about when there will be specificity or generality of transfer. Specificity, or limited transfer, occurs for tasks based primarily on procedural information, or skill, whereas generality, or robust transfer, occurs for tasks based primarily on declarative information, or facts. Thus, for skill learning, retention is strong but transfer is limited, whereas for fact learning, retention is poor but transfer is robust. In the last two lines of investigation that we present, we provide evidence supporting this working hypothesis.

Speeded Aiming

The first line of investigation (Healy, Wohldmann, Sutton, & Bourne, 2006) illustrates clearly the striking specificity of learning. This research involved a speeded aiming task in which subjects saw on a computer screen a clock face display with a central start position surrounded by a circle of digits, shown in Figure 1. A target digit was displayed above the start position, and subjects used a computer mouse to move a cursor from the start position to the location of the

digit around the circumference of the clock face. The task was made more difficult by reprogramming the computer mouse to introduce stimulus-response incompatibilities. Three reprogrammed mouse conditions were used: Either only horizontal movements were reversed (so when the mouse went left the cursor went right and vice versa), only vertical movements were reversed (so when the mouse went up the cursor went down and vice versa), or horizontal and vertical reversals were combined (so when the mouse went in any direction the cursor went in the opposite direction). Subjects were trained in one condition for 5 blocks of 80 trials and then returned 1 week later for testing for another 5 blocks of 80 trials in either the same or a different reprogrammed mouse condition.

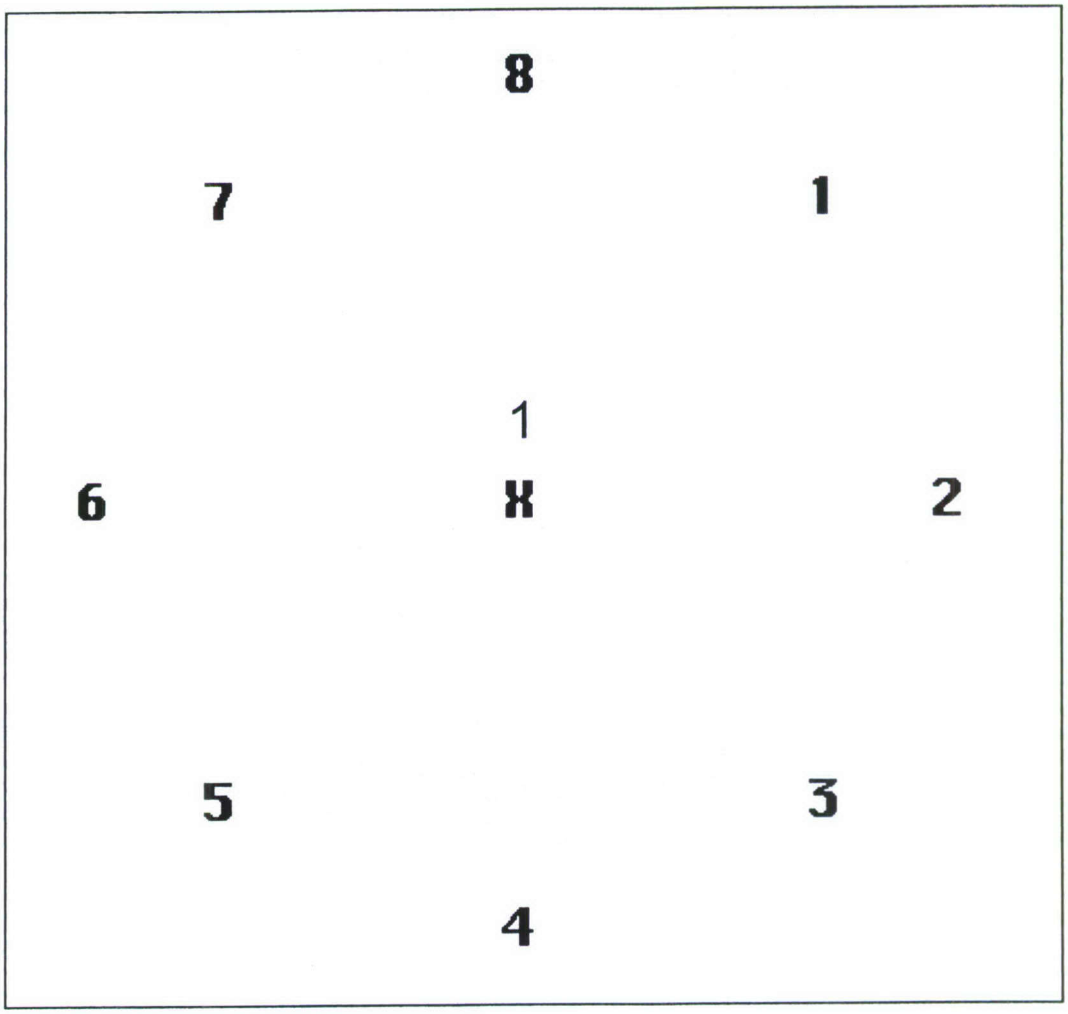

Figure 1. Clock face stimulus display.

The results are summarized in Figure 2 for the measure of movement time, which is the time to move from the start position to the target location. Comparisons of performance at the start and end of training showed a large decrease in movement time, demonstrating learning of this skill. Comparisons of performance at the end of training and the beginning of testing 1 week later for those subjects who were in the same reprogrammed mouse condition in both weeks also showed a small but significant decrease in movement time, reflecting perfect retention and

dissipation of fatigue across the 1-week delay. However, for those subjects who were in different reprogrammed mouse conditions in training and testing, there was actually a trend for movement time at the start of testing to increase relative to that at the start of training. Although subjects learned much during training, they could not transfer the skill they learned to training on a new condition a week later.

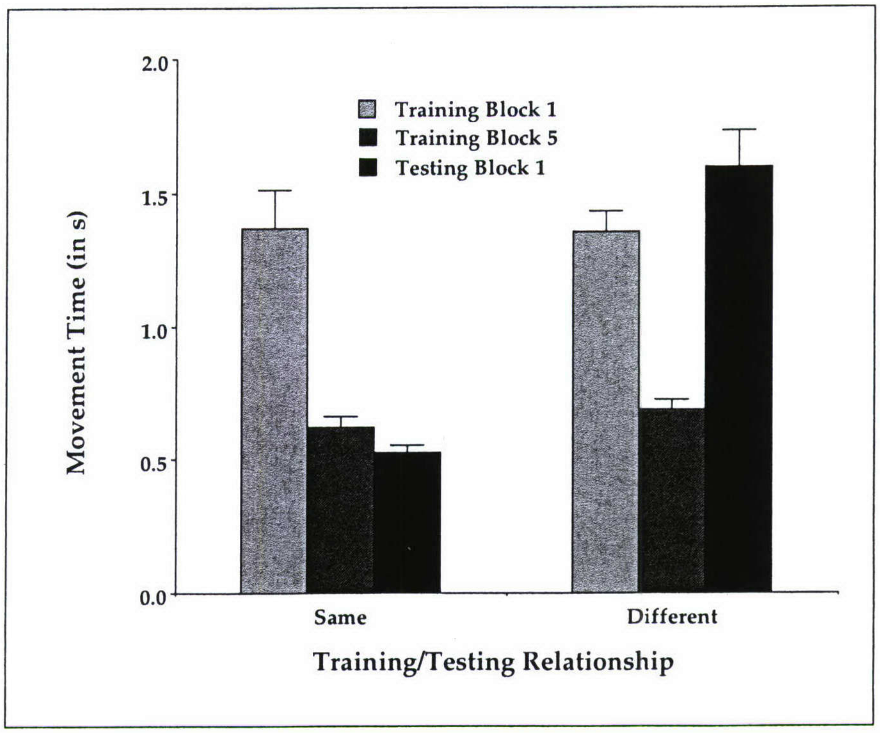

Figure 2. Response movement time (in s) as a function of the relationship between training and testing in reprogrammed mouse conditions for Training Block 1, Training Block 5, and Testing Block 1.

Note. Error bars represent positive standard errors of the mean. From Healy, Wohldmann, Sutton, and Bourne (2006, Experiment 1).

To understand the processes responsible for the severe specificity of training, we examined performance on the first block of training and on the first block of testing as a function of the specific reprogrammed mouse conditions employed for those subjects who were in different conditions in the 2 weeks. Figure 3 shows these results, where H represents the horizontal condition, V the vertical condition, and C the combined condition. Because the combined condition includes both types of reversals whereas the horizontal and vertical conditions include only a single reversal each, the first two sets of bars (labeled C/H and C/V) represent a change in the whole-part direction. Clearly there was interference in this case, with movement times at the start of testing slower than those at the start of training. Likewise, the bars labeled H/V and V/H represent a change in the part-part direction and show interference. The only bars showing some positive transfer, where movement time at the start of testing is somewhat faster than at the start of training, are labeled H/C and V/C and represent a change in the part/whole direction.

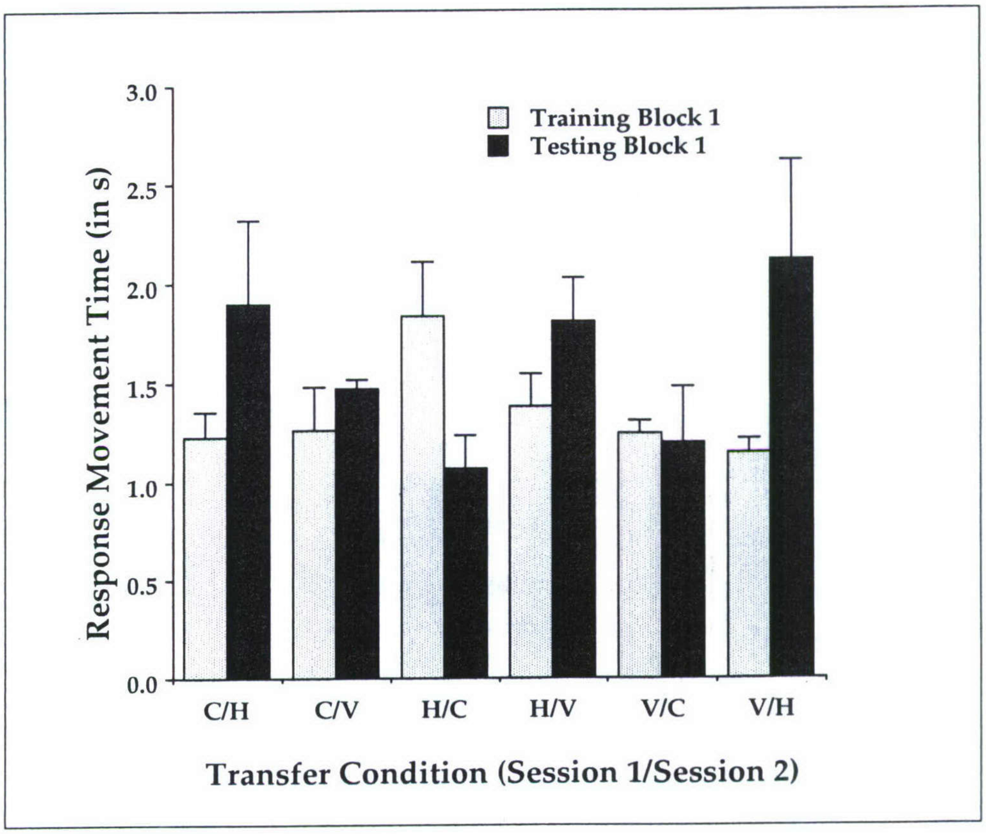

Figure 3. Mean movement time (in s) for the first block of training and the first block of testing as a function of transfer condition for only those participants who had a reprogrammed mouse in each session and switched reversal conditions across sessions.

Note. Error bars represent positive standard errors of the mean. C = combined, H = horizontal, V = vertical. From Healy, Wohldmann, Sutton, and Bourne (2006, Experiment 1).

Further insight into the underlying processes is derived from an examination of the type of movement made as a function of condition. Three different movement types are compared in Figure 4 as a function of reversal condition for performance at the start of training. Movement is either made to a target along the horizontal dimension (2 or 6), to a target along the vertical dimension (4 or 8), or to a target along one of the diagonals (1, 3, 5, or 7). It is interesting to note that subjects in the horizontal condition are actually faster for pure horizontal movements, which are reversed, than for pure vertical movements, which require no reversal. Similarly, subjects in the vertical condition are faster for pure vertical movements, which are reversed, than for pure horizontal movements, which require no reversal. Note also that for subjects in these two conditions, the movements are slowest along the diagonal. In contrast, diagonal movement is no slower than vertical or horizontal movement for subjects in the combined condition, and subjects in the combined condition are faster for movement along the diagonal than are subjects in the horizontal or vertical conditions.

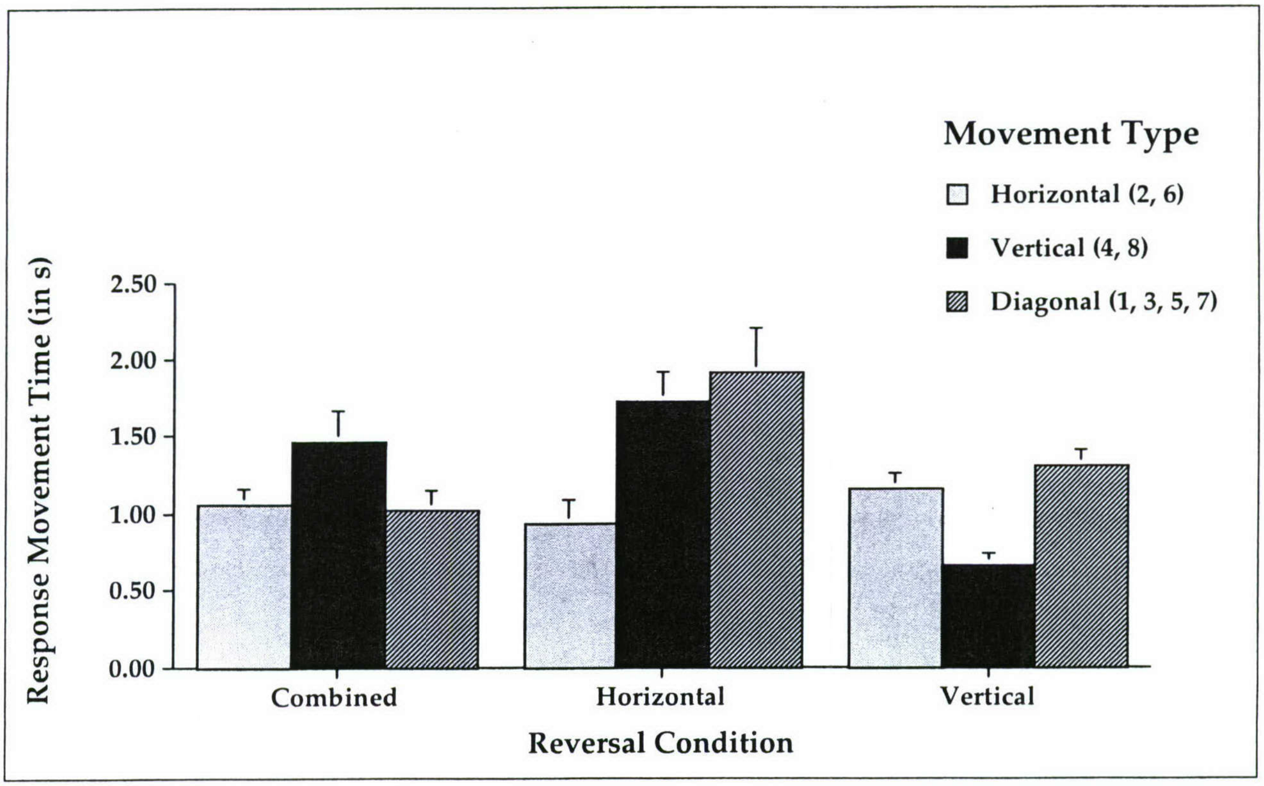

Figure 4. Mean movement time (in s) for Training Block 1 as a function of reversal condition and movement type.

Note. Error bars represent positive standard errors of the mean. From Healy, Wohldmann, Sutton, and Bourne (2006, Experiment 1).

We proposed a global inhibition hypothesis to explain these results. According to this hypothesis, whenever the mouse is reprogrammed, subjects apply a global inhibition strategy, inhibiting normal movements in *both* dimensions. When only one dimension is reversed, a further step is required to disinhibit responses on the non-reversed dimension. This strategy can explain why we saw positive part-whole transfer and negative or no transfer in the part-to-part or whole-to-part directions. For part-whole transfer, the global strategy applies directly to the transfer task, and there is no need for disinhibition during testing. In contrast, for part-part transfer, the inhibited and disinhibited dimensions must be exchanged between training and testing, and for whole-part transfer, one of the previously inhibited dimensions must be disinhibited during testing.

The global inhibition hypothesis also can explain why movement times are faster along a purely reversed dimension than along a dimension requiring no reversal in the horizontal and vertical conditions. According to the global inhibition hypothesis, inhibition is applied globally whenever some reversal is required, so that a further step is needed when only one dimension is reversed. Responses on the intact dimension must be disinhibited in these cases, causing slower responding on the non-reversed than on the reversed dimension and faster responding on diagonal movements in the combined condition (when no disinhibition is required) relative to the single reversal conditions (when disinhibition is required along one dimension).

To test the global inhibition hypothesis, we recently conducted a new pair of experiments (Healy, Wohldmann, & Bourne, 2007; Wohldmann, Healy, & Bourne, 2007b). In these experiments, subjects were always trained and tested for 200 trials in the horizontal condition. Also, only a 5-minute rest period separated training from testing, rather than a 1-week delay. Unlike the earlier experiment, in this new research, training involved only a subset of the target locations, whereas testing involved all locations. This method allowed us to examine transfer from trained target locations to untrained locations. Specifically, in Experiment 1 subjects were trained on two dimensions (either the pure horizontal or the pure vertical dimension and a diagonal), whereas in Experiment 2 subjects were trained on only a single dimension. Subjects in both experiments were tested on all dimensions. According to the global inhibition hypothesis, positive transfer to the untrained targets along the diagonal axis should be evident whenever training involves moving along a diagonal axis because both the trained and untrained targets along that axis demand that horizontal movements be inhibited and vertical movements be disinhibited in the same way. The global inhibition hypothesis thus would thus not predict any differences in transfer for trained and untrained movements. However, the strong degree of specificity found in earlier research suggests that the learned skill may depend on the particular target locations practiced as well as on such strategies as global inhibition. If specificity does apply in this case, then we should find faster movement times at test to the trained than to the untrained targets on the diagonal axes.

One possible way to overcome specificity of training and to promote transfer is to introduce variability into the practice routine (e.g., Schmidt & Bjork, 1992). These experiments also enabled us to evaluate this variability of practice hypothesis. For this evaluation, we examined test performance on diagonal targets and we compared Experiment 2, which involved training on only two targets, with Experiment 1, which involved more variable training on four targets. In both experiments testing was conducted on all eight targets. We broke down performance into sub-blocks of 16 trials (2 trials with each target). This analysis was limited to subjects who trained with one diagonal and tested with both diagonals; thus, all subjects in Experiment 1 but only half of the subjects in Experiment 2 were included. As shown in Figure 5, we found general task improvements across blocks of testing. Importantly, we also found that movement to old targets was significantly faster than movement to new targets, documenting specificity of training. Naturally, there was much more improvement across sub-blocks for new targets than for old ones. Also of interest with respect to the variability of practice hypothesis was the significant interaction between target type and experiment, reflecting a larger advantage for Experiment 1 relative to Experiment 2 (i.e., for training with four relative to training with two locations) with new targets than with old targets. Thus, variability of practice does seem to enhance transfer of training to new targets. In any event, finding specificity of training in this situation cannot be explained solely on the basis of the global inhibition hypothesis. Presumably subjects learned a strategy of inhibiting all normal movements but disinhibiting those along the vertical axis. In addition, however, they must have learned special, unique movement tactics to reach the targets along the trained diagonal, and those movement tactics could not be fully transferred to targets along the untrained diagonal.

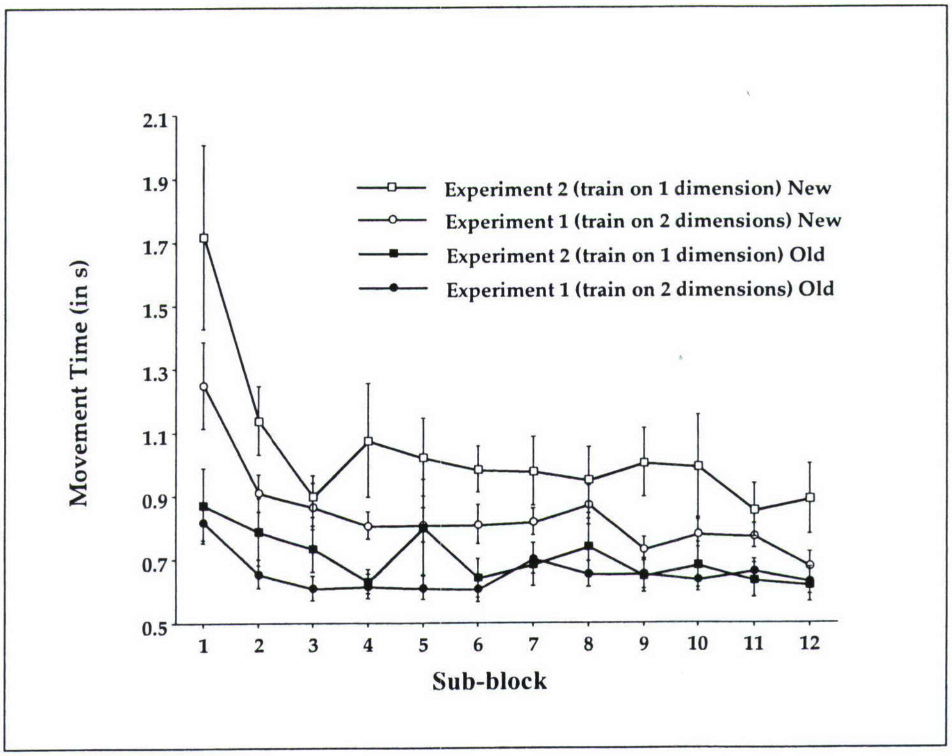

Figure 5. Mean movement time (in s) during testing for diagonal movements as a function of experiment, target type (old, new), and sub-block (1-12) in Experiments 1 and 2.

Note. Error bars represent positive and negative standard errors of the mean. From Wohldmann, Healy, and Bourne (2007b, Experiments 2 and 3).

Time Production

The high degree of specificity found for this task involving speeded aiming movements was surprising. However, perhaps even more surprising was the high degree of specificity we found in a second line of investigation (Healy, Wohldmann, Parker, & Bourne, 2005) because in this case we found a lack of transfer across different secondary, or background, tasks even when the primary, or foreground, task was held constant. In this research, subjects were trained to produce time intervals expressed in arbitrary units, with one unit equal to 783 ms. Subjects were not told how long a unit was, but they learned how to produce intervals by feedback on each response. For example, on a given trial, subjects may be told that after the beep they should estimate 32 units. They then pressed the space bar when they thought 32 units had elapsed. At that point they received feedback, such as "Your estimate was 29 units," and "The difference is -3 units."

Subjects practiced this task under one of two conditions. In the *no alphabet* condition, they performed no secondary task, whereas in the *alphabet* condition, they performed a secondary task in which they counted backwards through the alphabet by threes. When subjects reached the beginning of the alphabet, they were to revert to the end of the alphabet and continue from that point. For example, if they were given the letter cue *s*, they were to say *s*, *p*, *m*, *j*, *g*, *d*, *a*, *x*. This is clearly a difficult secondary task. Subjects were trained in one condition for six

blocks of six trials, and then they returned 1 week later for six blocks of six trials of testing in a condition that was either the same as or different from the training condition. We used two measures of performance for this task. The *proportional absolute error* is the absolute (or unsigned) difference between the produced interval and the specified interval divided by the specified interval. This index provides a normalized assessment of error magnitude. In contrast, the *proportional relative error* is the signed difference between the produced interval and the specified interval divided by the specified interval. The proportional relative error is just like the proportional absolute error but uses signed differences instead of absolute differences. It provides an index of response bias. When the produced interval is longer than the specified interval, there is positive bias by this index, whereas when the produced interval is shorter than the specified interval, there is negative bias. Figure 6 allows us to compare performance in terms of proportional absolute error at the start and end of training, and shows that subjects improved their skill of time production during training. Also, comparing performance at the end of training and the start of testing when subjects were in the same condition in training and testing reveals no decrement in performance, thereby showing perfect skill retention. However, when subjects were in different conditions in training and testing, performance did suffer across the 1-week delay, so that, in fact, performance at the start of testing was comparable to that at the start of training. Thus, again, as with the clock face task involving speeded aiming movements, there was perfect retention but no transfer of this time production skill from one condition to another, even though the required skill did not change; instead the only change was in the secondary, background task.

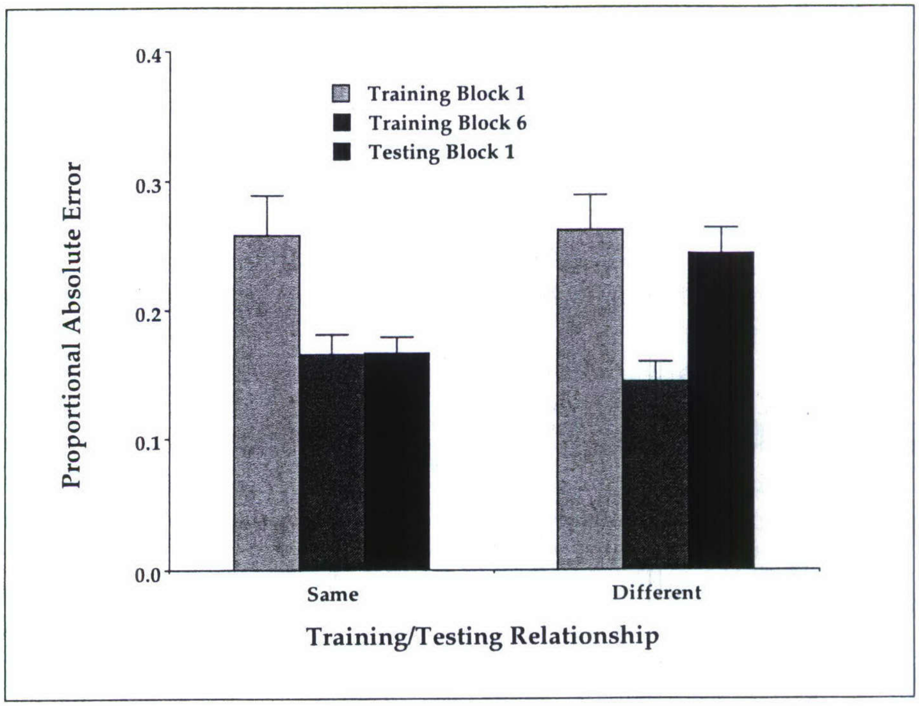

Figure 6. Proportional absolute error as a function of the relationship between training and testing in time production conditions for Training Blocks 1 and 6, and Testing Block 1.

Note. Error bars represent positive standard errors of the mean. From Healy, Wohldmann, Parker, and Bourne (2005, Experiment 2).

We explained the high degree of specificity in this case by assuming that what subjects learned about time production incorporated the secondary alphabet task so if the secondary task was changed (even if it was simply removed) the subjects could not benefit at all from training. More generally, we proposed a functional task principle, according to which secondary task requirements are often integrated with primary task requirements during learning, resulting in the acquisition of a single functional task rather than two separate tasks. For example, subjects might learn, not that a given unit is about 3/4 s, but rather how many letters in the alphabet task they should count before a given number of units has elapsed.

We recently completed another experiment aimed to provide a test of this functional task principle (Wohdlmann, Healy, & Bourne, 2007a). One basic prediction from the functional task principle is that if performance on the secondary task is affected by task difficulty, then there should be a corresponding, congruent change in performance on the primary task, and there should be no tradeoff in performing the two tasks. Specifically, subjects were trained to produce three different intervals of time while performing one of two versions of the alphabet counting task. In the *fixed* alphabet condition, the starting letter was always the same (*m*), whereas in the *random* alphabet condition, the starting letter varied randomly across trials. In this experiment, unlike the last one, subjects typed the letters for the secondary task (they did not just say them aloud), so we could easily measure performance on the secondary task as well as on the primary task. After six blocks of training, each including six trials, subjects engaged in an irrelevant letter detection task, which took about 15 minutes to complete. Then subjects were tested under the same conditions for another six blocks of six trials.

We expected accuracy on the time production task to improve across blocks of trials and, based on the findings from our earlier experiment, we also expected perfect retention for the time production task between training and testing. Of greater interest was the relationship between the changes in performance on the two tasks. According to the functional task principle, we expected performance on both the time production and alphabet-counting task to be better when the starting letter was fixed rather than randomly changing from trial to trial. This expectation was based on the assumption that subjects develop a strategy for time production that is based on the number of letters counted for a given interval. That strategy should be easier to employ with a fixed starting letter because subjects could simply learn on what letter to stop for each of the three time intervals tested when the starting letter is fixed but they would have to learn many more rules of that type when the starting letter is random. However, fixing the starting letter might also have an adverse effect on time production if that condition promotes a continuous increase across trials in the number of letters typed for a given time interval. Under those circumstances, there might be a negative bias in time production so that the intervals produced would be shorter than those prescribed. Finding these specific patterns for the comparison of the fixed versus random conditions would be evidence for the functional task principle and inconsistent with alternative theories based on attentional bottlenecks or resource limitations when performing two tasks simultaneously (e.g., Broadbent, 1958; Gopher, Brickner, & Navon, 1982).

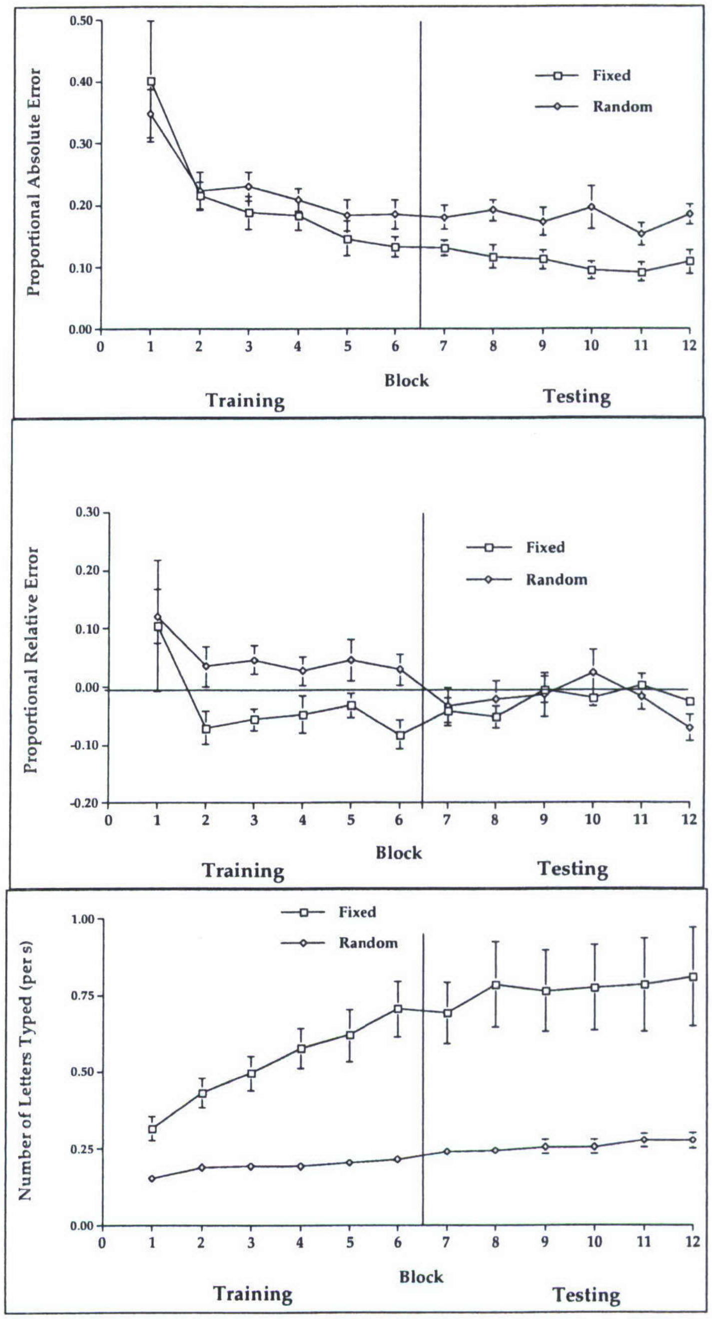

Figure 7. Proportional absolute error for the primary task (top panel), proportional relative error for the primary task (middle panel), and number of letters typed per second on the secondary task (bottom panel) as a function of session half, block, and alphabet condition.

Note. Error bars represent positive and negative standard errors of the mean. From Wohldmann, Healy, and Bourne (2007a, Experiment 1).

The results are summarized in Figure 7, which has three panels. The top panel shows primary task performance in terms of proportional absolute error, the normalized index of error

magnitude. The middle panel shows primary task performance in terms of proportional relative error, the index of response bias. The bottom panel shows performance on the secondary task as a function of the number of letters typed per second. We found that on the secondary task subjects typed many more letters per second when the starting letter was fixed than when it was random. Also, the number of letters typed per second showed a large, steady increase across blocks in training when the starting letter was fixed but a minimal increase when the starting letter was random. Despite the changes in performance on the secondary task with the fixed starting letter, performance on the primary time production task showed significantly less proportional absolute error with the fixed than with the random starting letter, especially during testing. This result can be explained by assuming that subjects counted the number of letters said in the secondary task as a means to gauge how much time had elapsed. However, if subjects used a strategy that required them to stop after a certain number of letters were counted, then they would stop earlier whenever they were able to go through the alphabet more quickly. This strategy would lead the subjects to produce shorter time intervals than those prescribed when the number of letters typed per second increased, resulting in a negative bias in the subjects' time judgments. Thus, subjects should show a negative bias on the primary task in terms of proportional relative error whenever the number of letters said per second on the secondary task increased. Consistent with this reasoning is the finding for the fixed condition that as the number of letters said per second increased during training there was a negative bias. During testing, the number of letters said per second did not change markedly, and there was little bias.

The results of this experiment provide clear-cut evidence that improvements in the secondary task do have an influence on the level of improvement in the primary task, in accordance with the functional task principle. Changes in secondary task performance (an increase in the number of letters said per second across blocks in the fixed alphabet condition) were generally consistent with changes in primary task performance (negative bias across blocks during training and lower errors across blocks during testing in the fixed alphabet condition). Thus, it did not appear that subjects traded off one task for the other; rather, they seemed to develop a strategy that combined the requirements of the two tasks.

Navigation

Although the functional task principle can help us understand the high degree of specificity or lack of transfer we found for the time production task, this principle does little to help us understand the specificity we recently found in an experiment in another line of investigation. We have been studying a situation meant to mimic communication between air traffic controllers and flight crews, which involves giving and receiving navigation instructions (Barshi & Healy, 2002; Schneider, Healy, & Barshi, 2004). Flight crews often make errors resulting from this communication, and these errors can lead to serious accidents. Thus, our research has the eventual goal to determine ways to reduce such critical errors.

In our task, subjects receive messages instructing them to make movements in a space shown on the computer screen. Figure 8 shows on the left the three-dimensional space that is depicted on the computer screen by the grid of matrices shown on the right. In our experimental task, the messages vary in length from one to six commands. A sample message with three commands (Length 3) is: Left 2 squares, down 2 levels, forward 1 step. Subjects in our

experiments typically hear the commands; they do not see them. Their task is to repeat back immediately the commands heard and then to follow them by clicking, with a computer mouse, on the appropriate squares of the grid. In the experiment just completed, we were interested in how best to train subjects to perform this task (Schneider, Healy, Barshi, & Bourne, 2007). We wondered whether it would be better to provide *easy* training (restricted to short message lengths 1-3), *hard* training (restricted to long message lengths 4-6), or *mixed* training (with all six message lengths 1-6). Testing in all conditions was in the mixed format involving all six message lengths. We used an all-or-none scoring system to measure performance in terms of the proportion of correct responses. All required clicks had to be made in order for a given trial to be scored as correct.

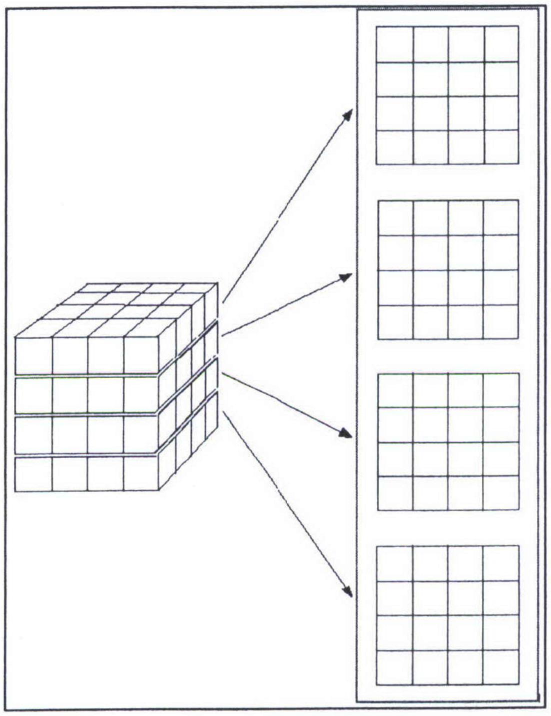

Figure 8. Graphic showing the three-dimensional space (left) depicted on the computer screen by the grid of matrices shown on the right.

Note. From Schneider, Healy, Barshi, and Bourne (2007).

This experiment allowed us to test again for the advantage of variability of practice and also to see whether specificity of training applies in this paradigm. On the basis of the predicted advantage for variability of practice, subjects should do better at test with the more variable mixed training than with either easy or hard training. Because testing involves all six lengths, finding an advantage at test for mixed training also would be consistent with specificity of training. Further, on the basis of specificity of training, at test with short message lengths subjects should do better with easy than with hard training, but with long message lengths they should do better with hard than with easy training. Thus, there should be an interaction between training condition and message length.

As shown in Figure 9, there was evidence for both these hypothesized patterns of results at test. First note that in this experiment, as in others we have conducted with this paradigm, there is a huge effect of message length, with performance dropping off markedly when the number of commands per message increases from three to four. Also note that, as expected on the basis of the predicted advantage for variability of practice, performance was best in the mixed condition at almost all message lengths. This finding is consistent with specificity of training, as is the observed interaction between training condition and message length, and this interaction was significant even when considering only the easy and hard training conditions. It is clear that subjects given easy training performed better than did subjects given hard training on the short message lengths (1-3), whereas subjects given hard training performed better than did subjects given easy training on the long message lengths (4-6). Therefore, test performance on a given set of message lengths was better for subjects trained with those lengths than for subjects trained with the other set of lengths. Thus, we have shown that training is specific even to the length of messages that need to be understood and executed.

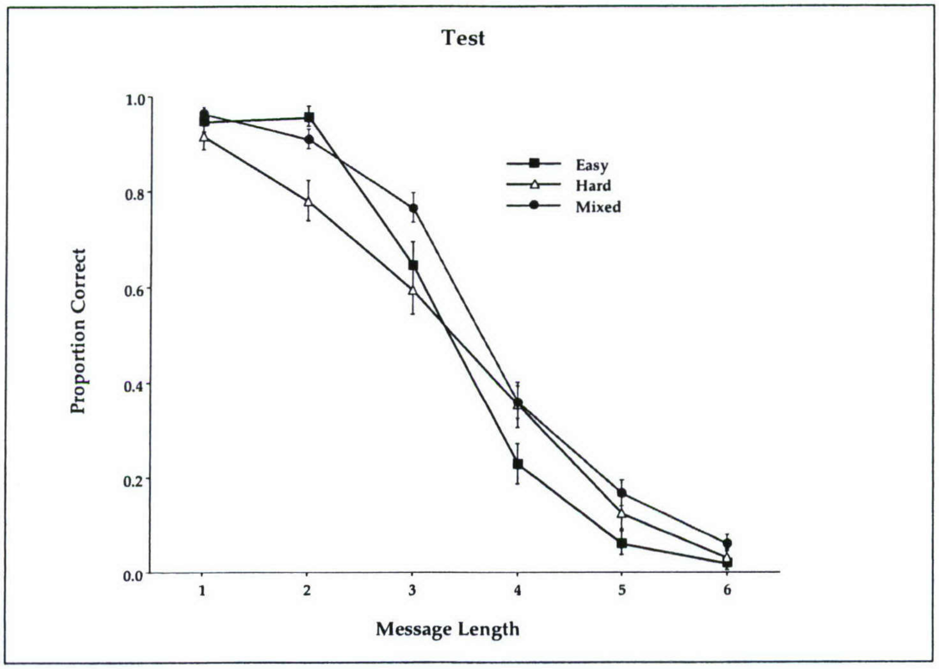

Figure 9. Proportion correct as a function of training condition and message length.

Note. Error bars represent positive standard errors of the mean. From Schneider, Healy, Barshi, and Bourne (2007).

Fact Learning

We have reviewed studies documenting the first part of our working hypothesis (Healy, 2007) by showing limited transfer for learning skills. Now we turn to studies documenting the second part of our working hypothesis by showing robust transfer of knowledge for learning facts. These studies demonstrate that learning is facilitated whenever pre-existing knowledge can be employed as a mediator in the process of acquisition. The first research of this type we will review shows just how powerful transfer can be for such a task. In this work (Kole & Healy, 2007), subjects started with a listing procedure in which they gave us the names of 12 individuals, 6 male and 6 female, whom they know very well, such as their friends or relatives. Later, all subjects in each of three groups learned a total of 144 fictitious facts, consisting of 12 facts about each of 12 individuals. Subjects in the *high knowledge* group learned facts about the 12 well-known individuals whom they had listed earlier. Subjects in the *low knowledge* group learned facts about 12 relatively unfamiliar individuals; the names used in this group were drawn from the fields of professional beach volleyball and off-Broadway plays. Subjects in the *mediated knowledge* group also learned facts about 12 unfamiliar individuals, the same set as in the low knowledge group, but these subjects associated each unfamiliar individual with a well-known individual. Specifically, after the listing procedure, subjects in the mediated knowledge group completed association training, during which they were trained to associate the 12 people whom they had just listed with the 12 unfamiliar individuals whom they would learn about later, in a paired-associate task, to a criterion of three consecutive efforts test cycles with perfect accuracy.

The 144 facts subjects learned were presented as sentences in the format of name, verb phrase, and fact category exemplar. Each was a unique, fabricated fact. There were 12 fact categories used, all commonly known attributes about people, such as their favorite author, the type of car they drive, what sport they play, and their favorite food. The fact category exemplars were each one word in length, and were common instances of the category. So, for example, if Alice Healy was a subject in the high knowledge group, she might learn that her daughter Charlotte Healy likes to play lacrosse (which is certainly not true). If she was in the low knowledge group, she might learn that Nancy Reno likes to play lacrosse. If she was in the mediated knowledge group, she also might learn that Nancy Reno likes to play lacrosse, but she would be encouraged to think of Charlotte Healy while learning this fact. These 144 facts were presented in a learning phase, which consisted of a study-test procedure. Each sentence was presented individually, in blocks of 12 sentences. After each block there was a cued recall test over the 12 facts in the block. The presentation and testing of the entire set of 144 sentences constituted one learning round, and subjects completed three learning rounds altogether. After the three learning rounds, subjects were given a final cued recall test, during which they were examined on all 144 facts without further study.

The results for the final test are summarized in Figure 10. There was a significant main effect of knowledge group, as evidenced by the fact that accuracy was highest for the high knowledge group and lowest for the low knowledge group, with the mediated knowledge group in between. The difference between the mediated and low knowledge groups was significant. We conclude that prior knowledge can aid in the learning of new information. Furthermore, the

facilitative effect of using prior knowledge to learn new information was found even when the prior knowledge played only a mediating function, rather than a direct function.

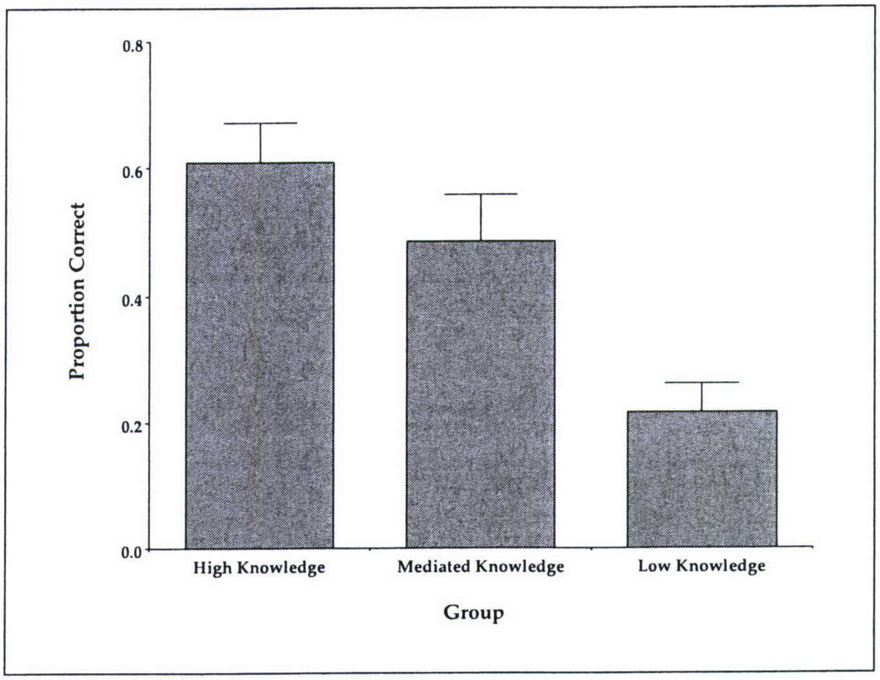

Figure 10. Proportion correct during the final test phase as a function of knowledge group in Experiment 1.

Note. Error bars represent positive standard errors of the mean. From Kole and Healy (2007, Experiment 1).

In this experiment, prior knowledge was used to learn new information conceptually related to the prior knowledge. That is, subjects used their prior knowledge of people to learn about people. Thus, the question arises whether the facilitative effect of mediated learning depends on whether the information to be learned is conceptually related to the prior knowledge. For example, conceptually unrelated mediated learning might be using one's knowledge of people to learn about countries. Experiment 2 addressed this question.

In Experiment 2, the domain, or type of facts subjects learned was manipulated. As in the previous experiments, subjects learned a total of 144 facts. In the person domain, subjects learned fictitious facts about 12 unfamiliar individuals, and in the country domain, subjects learned true facts about 12 relatively unfamiliar countries. Knowledge group was also varied, with subjects in the low knowledge group learning facts about these unfamiliar individuals or countries, and those in the mediated knowledge group learning the same facts while associating the unfamiliar countries or individuals with well-known individuals. The stimuli employed were constructed such that the fact category exemplars learned were the same across conditions. So, for example, in the person domain, subjects learned that Linda Hanley's monthly income is $2200, and in the country domain, they learned that Singapore's gross domestic product per capita is $2200. The method was largely the same as in the previous experiment: Again, there was a listing procedure, during which all subjects listed the names of 12 individuals whom they

know well. Subjects in the mediated knowledge groups only completed association training, whereby they were trained to associate the 12 unfamiliar individuals or countries with the names of the 12 familiar individuals. The number of association training rounds was fixed at eight. Then the learning phase and test phase proceeded as before, except that subjects were tested with a cued recognition test. Specifically, subjects were provided with a sheet giving all the possible answers for each of the 12 fact categories.

As shown in Figure 11, the proportion of correct responses on the test for mediated knowledge subjects, who had been trained to associate the unfamiliar items with familiar individuals, was overall more than twice as high as that for low knowledge subjects, who had not received such association training. This advantage for mediated learning was just about as large for learning facts about countries (which are conceptually unrelated to the familiar individuals) as for learning facts about other people. Also, prior knowledge about familiar individuals aided learning facts about unfamiliar individuals even though the facts were unlikely to be true about the familiar individuals with whom they were associated. We explained these results in terms of a mental model approach, following Johnson-Laird (1983), Radvansky and Zacks (1991), and others. Specifically, prior knowledge about well-known individuals can be viewed as mental models or integrated representations, and the new facts learned can be incorporated easily within these well-established, integrated representations, thereby minimizing interference both from other facts learned about the same individuals and from similar facts learned about other individuals.

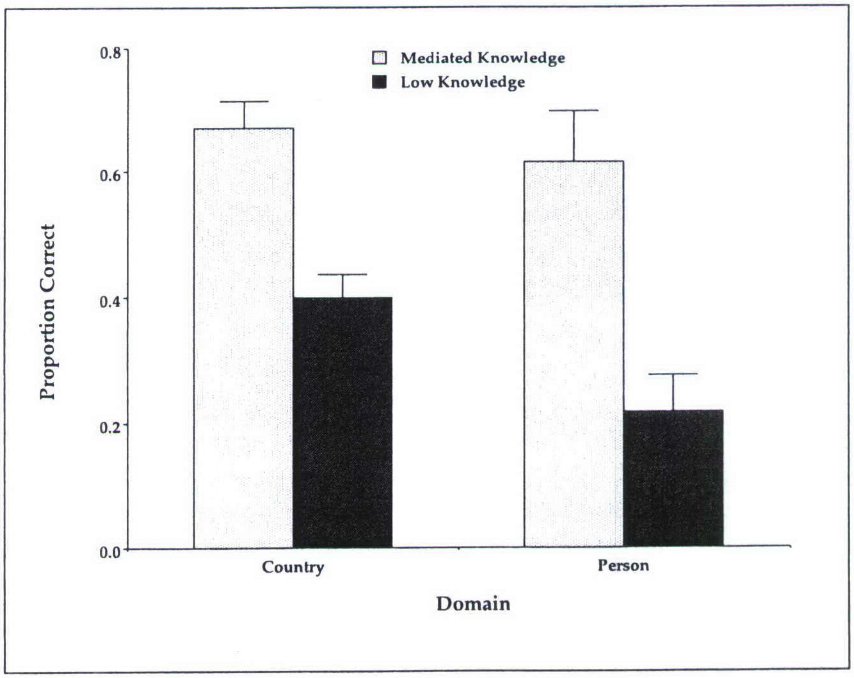

Figure 11. Proportion correct during the final test phase as a function of knowledge group and domain in Experiment 2.

Note. Error bars represent positive standard errors of the mean. From Kole and Healy (2007, Experiment 3).

Memory for Order Information

More recently, we conducted an experiment in another series in which we also found a surprising ability to transfer information from prior knowledge to learn a new set of declarative facts (Healy, Shea, Kole, & Cunningham, in press). In this line of investigation we are attempting to understand the processes underlying the bow-shaped serial position function evident in studies of memory for order information. Earlier experiments in this series examined memory for information either learned outside the laboratory or learned in a single list presentation in the laboratory. In our new experiment, we used a novel training procedure in which list items were presented two to four times each. With this procedure we compared the relative contributions of two variables postulated to affect the serial position functions: familiarity of the items and distinctiveness of the positions. Another novel aspect of the present research was that it tested memory for spatial order, rather than memory for temporal sequence. Specifically, each item occurred individually on a different numbered line in a vertical array. Thus, unlike earlier studies, position distinctiveness was not confounded with temporal distinctiveness.

Specifically, subjects learned the spatial positions of a list of 20 names. Half of these names were familiar names of friends and relatives, which we obtained using the same listing procedure we had employed in our last effort. The remaining half of the names were unfamiliar ones, taken from the set of professional beach volleyball players and actors from off-Broadway plays that we had used in our other research. We compared two different familiarity conditions. In both of these conditions, the familiar and unfamiliar names alternated in pairs. In the *familiar first* condition, familiar names were assigned to the first two list positions, unfamiliar names were assigned to the next two list positions, and so on throughout the list of 20 names. The opposite assignment of names to list positions was used for the *unfamiliar first* condition. Following other procedures we developed earlier, at test subjects were asked to reconstruct the order of only a 12-name subset of the full list of 20 names. There were three serial position conditions that differed in terms of the 12-name subset of positions used for the reconstruction of order task: 1-12, 5-16, and 9-20. Subjects were given the 12 names in the subset, listed alphabetically, and had to type 1 of the 12 possible position numbers beside each name.

As shown in Figure 12, there was a striking effect of serial position condition on the shape of the serial position function, with a strong primacy effect only for the 1-12 condition and a strong recency effect only for the 9-20 condition. This pattern of results suggests that distinctiveness of the absolute list positions, rather than that of the relative list positions, contributes to the bow-shaped serial position functions. Most interesting are the effects of familiarity condition shown in Figure 13. There were huge effects of familiarity. For the familiar first condition, the first two positions are much higher than the next two positions, and so on throughout the list, whereas the opposite pattern is found for the unfamiliar first condition. This scalloped pattern reveals that reconstruction of order performance is almost twice as high on average for the familiar names as for the unfamiliar names. These effects of item familiarity vividly illustrate that subjects can use their prior knowledge about items to learn the spatial locations of the items. Again a mental model account of memory can be used to explain the results, according to which prior knowledge of individuals can be viewed as integrated representations, and new facts about the individuals (in this case those pertaining to spatial

location information) can be incorporated within these integrated representations. Thus, for example, Alice Healy could use her extensive knowledge of Charlotte Healy to elaborately encode her position as Number 5.

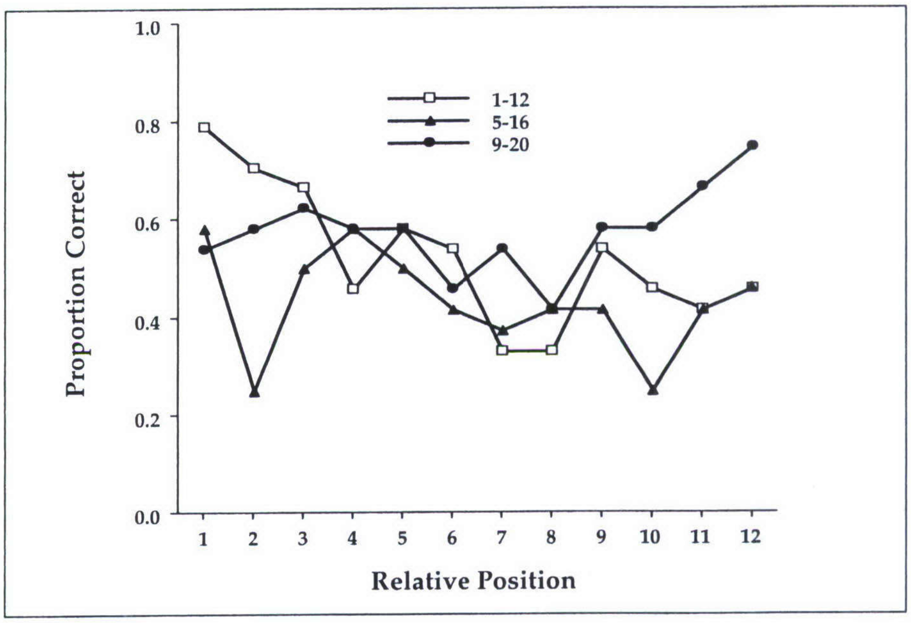

Figure 12. Proportion of correct responses in the reconstruction of order task as a function of relative serial position and serial position condition.

Note. From Healy, Shea, Kole, and Cunningham (in press, Experiment 3).

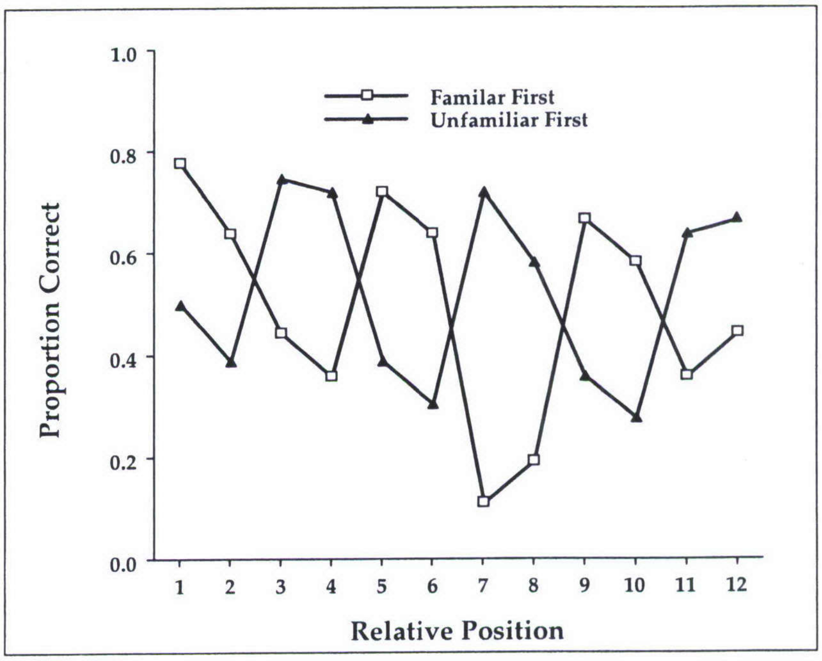

Figure 13. Proportion of correct responses in the reconstruction of order task as a function of relative serial position and familiarity condition.

Note. From Healy, Shea, Kole, and Cunningham (in press, Experiment 3).

Summary, Conclusions, and Applications

These last two lines of investigation show that we can indeed use prior learning to aid in new learning of declarative information. These dramatic benefits of prior learning for declarative information stand in contrast to the surprising lack of benefits for procedural information found in the first three lines of investigation. Together these studies support our working hypothesis that expands on the earlier principles and models of encoding specificity (Tulving & Thomson, 1973), transfer appropriate processing (McDaniel et al., 1978; Morris et al., 1977; Roediger et al., 1989), procedural reinstatement (Healy & Bourne, 1995; Healy, Wohldmann, & Bourne, 2005), and identical elements (Rickard, 2005; Rickard & Bourne, 1995; Singley & Anderson, 1989; Thorndike, 1906). These principles and models can account for the effects of training specificity that were observed, but they cannot explain without expansion when generalizability will occur instead of specificity. According to our expanded working hypothesis (Healy, 2007), there is specificity (limited transfer) for tasks based primarily on procedural information, or skill, whereas there is generality (robust transfer) for tasks based primarily on declarative information, or facts. Thus, we propose that for skill learning, retention is high but transfer is low; in contrast, for fact learning, retention is low but transfer is high. In our present research we are conducting further tests of this working hypothesis, and we are trying to develop new ways, similar to variability of practice, to overcome the limited transfer of procedural information and new ways, similar to the use of mediated learning, to leverage the robust transfer for declarative information.

The research summarized in this report sets the stage for significant developments in applied research. To illustrate this potential symbiosis between basic and applied research, we give two brief examples. First, our research has demonstrated a high degree of specificity from training to subsequent application. In fact, we have shown that training is specific even to the length of messages that need to be understood and executed in the navigation situation (Schneider et al., 2007). Test performance was best following training with all possible message lengths. These findings have crucial implications for military training because instructors may assume that teaching a particular task through a limited number of examples will generalize fully to an entire domain even when the examples differ in a fundamental respect (e.g., length) from the test situations. However, our findings imply that to be effective, training should incorporate a full range of examples on critical task dimensions. Although the tasks used in our research are often components of military tasks, they are not the real military tasks currently being trained in the Army. We hope that applied research units of ARI will be interested in testing whether the principles we have developed apply in more realistic tasks and whether the methods we hope to develop for overcoming the problem of training specificity could be adapted to improve military training. For example, our findings of specificity imply that in military training using simulators, the simulator environment needs to match as closely as practically feasible the field context in which the trained military action will take place. Obviously, a simulator cannot match all aspects of the field; it would, thus, be imperative for applied research to determine which aspects are crucial in this regard.

Second, we have shown that prior knowledge can be used to enhance the learning of spatial position information (Healy et al., in press). We are presently conducting research to explore the boundaries of effective prior knowledge utilization. The use of prior knowledge also

needs to be examined in the context of real military tasks, and we hope that ARI applied research units will be willing and able to take on this examination. For example, in a field situation, Soldiers might need to learn facts about new places that they will encounter, including their location, their strategic importance, their size or scope, and the activities that take place there. Our findings imply that associating each place with some item from a familiar domain of knowledge, such as a friend or family member, can help Soldiers learn the necessary information quickly and retain it better. However, such a technique needs to be evaluated in a context closer than the laboratory setting to that actually encountered in the field before it should be applied wholesale.

References

Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.

Barshi, I., & Healy, A. F. (2002). The effects of mental representation on performance in a navigation task. *Memory & Cognition, 30,* 1189-1203.

Broadbent, D. E. (1958). *Perception and communication*. London: Pergamon Press.

Gopher, D., Brickner, M., & Navon, D. (1982). Different difficulty manipulations interact differently with task emphasis: Evidence for multiple resources. *Journal of Experimental Psychology, Human Perception and Performance*, *8,* 146-157.

Harlow, H. F. (1949). The formation of learning sets. *Psychological Review, 56,* 51-65.

Healy, A. F. (2007). Transfer: Specificity and generality. In H. L. Roediger, III, Y. Dudai, & S. M. Fitzpatrick (Eds.), *Science of memory: Concepts* (pp. 271-275). New York: Oxford University Press.

Healy, A. F., & Bourne, L. E., Jr. (Eds.). (1995). *Learning and memory of knowledge and skills: Durability and specificity*. Thousand Oaks, CA: Sage.

Healy, A. F., Shea, K. M., Kole, J. A., & Cunningham, T. F. (in press). Position distinctiveness, item familiarity, and presentation frequency affect reconstruction of order in immediate episodic memory. *Journal of Memory and Language*.

Healy, A. F., Wohldmann, E. L., & Bourne, L. E., Jr. (2005). The procedural reinstatement principle: Studies on training, retention, and transfer. In A. F. Healy (Ed.), *Experimental cognitive psychology and its applications* (pp. 59-71). Washington, DC: American Psychological Association.

Healy, A. F., Wohldmann, E. L., & Bourne, L. E., Jr. (2007, July). *Training specificity and global inhibition in speeded aiming movements*. Paper presented at the joint meeting of the Experimental Psychology Society and the Psychonomic Society, Edinburgh, Scotland.

Healy, A. F., Wohldmann, E. L., Parker, J. T., & Bourne, L. E., Jr. (2005). Skill training, retention, and transfer: The effects of a concurrent secondary task. *Memory & Cognition, 33,* 1457-1471.

Healy, A. F., Wohldmann, E. L., Sutton, E. M., & Bourne, L. E., Jr. (2006). Specificity effects in training and transfer of speeded responses. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32,* 534-546.

Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness.* Cambridge, MA: Harvard University Press.

Kole, J. A., & Healy, A. F. (2007). Using prior knowledge to minimize interference when learning large amounts of information. *Memory & Cognition, 35,* 124-137.

Kolers, P. A., & Roediger, H. L., III (1984). Procedures of mind. *Journal of Verbal Learning and Verbal Behavior, 23,* 425-449.

McDaniel, M. A., Friedman, A., & Bourne, L. E. (1978). Remembering the levels of information in words. *Memory & Cognition, 6,* 156-164.

Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior, 16,* 519-533.

Radvansky, G. A. & Zacks, R. T. (1991). Mental models and the fan effect. *Journal of Experimental Psychology: Learning. Memory, and Cognition, 17,* 940-953.

Rickard, T. C. (2005). A revised identical elements model of arithmetic fact representation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 250-257.

Rickard, T. C., & Bourne, L. E., Jr. (1995). An identical-elements model of basic arithmetic skills. In A. F. Healy & L. E. Bourne, Jr. (Eds.), *Learning and memory of knowledge and skills: Durability and specificity* (pp. 255-281). Thousand Oaks, CA: Sage.

Roediger, H. L., III, Weldon, M. S., & Challis, B. H. (1989). Explaining dissociations between implicit and explicit measures of retention: A processing account. In H. L. Roediger, III, & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 3-41). Hillsdale, NJ: Erlbaum.

Schmidt, R. A., & Bjork, R. A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychological Science, 3,* 207-217.

Schneider, V. I., Healy, A. F., & Barshi, I. (2004). Effects of instruction modality and readback on accuracy in following navigation commands. *Journal of Experimental Psychology: Applied*, 10, 245-257.

Schneider, V. I., Healy, A. F., Barshi, I., & Bourne, L. E., Jr. (2007, November). *Effects of difficulty, specificity, and variability on training to follow navigation instructions*. Poster to be presented at the 48th Annual Meeting of the Psychonomic Society, Long Beach, CA.

Singley, M. K., & Anderson, J. R. (1989). *The transfer of cognitive skill.* Cambridge, MA: Harvard University Press.

Thorndike, E. L. (1906). *The principles of teaching: Based on psychology*. New York: A. G. Seiler.

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80,* 352-373.

Wohldmann, E. L., Healy, A. F., & Bourne, L. E., Jr. (2007a). *Task integration under conditions of divided attention*. Manuscript submitted for publication.

Wohldmann, E. L., Healy, A. F., & Bourne, L. E., Jr. (2007b). *Global inhibition and midcourse corrections in speeded aiming*. Manuscript submitted for publication.